

ABSTRACT

A system (1) generates an output indicating scores for the extent of matching of pairs of data records. Thresholds may be set for the scores for decision-making or human review. A vector extraction module (12) measures similarity of pairs of fields in a pair of records to generate a vector. The vector is then processed to generate a score for the record pair.